

Tier-Aware Resource Allocation in OFDMA Macrocell-Small Cell Networks

Amr Abdelnasser, Ekram Hossain, and Dong In Kim

Abstract—We present a joint sub-channel and power allocation framework for downlink transmission an orthogonal frequency-division multiple access (OFDMA)-based cellular network composed of a macrocell overlaid by small cells. In this framework, the resource allocation (RA) problems for both the macrocell and small cells are formulated as optimization problems. For the macrocell, we formulate an RA problem that is aware of the existence of the small cell tier. In this problem, the macrocell performs RA to satisfy the data rate requirements of macro user equipments (MUEs) while maximizing the tolerable interference from the small cell tier on its allocated sub-channels. Although the RA problem for the macrocell is shown to be a mixed integer nonlinear problem (MINLP), we prove that the macrocell can solve another alternate optimization problem that will yield the optimal solution with reduced complexity. For the small cells, following the same idea of tier-awareness, we formulate an optimization problem that accounts for both RA and admission control (AC) and aims at maximizing the number of admitted users while simultaneously minimizing the consumed bandwidth. Similar to the macrocell optimization problem, the small cell problem is shown to be a MINLP. We obtain an upper bound on the optimal solution with reduced complexity through convex relaxation. In addition, we employ the dual decomposition technique to have a distributed solution for the small cell tier. Numerical results confirm the performance gains of our proposed RA formulation for the macrocell over the traditional resource allocation based on minimizing the transmission power. Besides, it is shown that the formulation based on convex relaxation yields a similar behavior to the MINLP formulation. Also, the distributed solution converges to the same solution obtained by solving the corresponding convex optimization problem in a centralized fashion.

Keywords: Base station densification, small cells, OFDMA, downlink resource allocation, sub-channel and power allocation, admission control, convex optimization, dual decomposition.

I. INTRODUCTION

As more and more customers subscribe to mobile broadband services, there is a tremendous growth in the demand for mobile broadband communications, together with the increased requirements for higher data rates, lower latencies and enhanced quality-of-service (QoS). Fueled by the popularity of smartphones and tablets with powerful multimedia capabilities, services and applications, it is anticipated that by 2020, the existing wireless systems will not be able to accommodate the expected 1000-fold increase in total mobile broadband data [2]. Therefore, 5G cellular technologies are being sought. Of

the several enabling technologies for 5G to handle the expected traffic demand, base station (BS) densification is considered as one of the most promising solutions [3].

BS densification involves the deployment of a large number of low power BSs. This decreases the load per BS and leads to a better link between a user equipment (UE) and its serving BS owing to a smaller distance between them [4]. This BS densification also creates a multi-tier network of nodes with different transmit powers, coverage areas and loads. Two issues arise in such a dense multi-tier network. The first issue is that the resource allocation (RA) in one tier cannot be done in isolation of the resource allocation in another tier. In other words, one tier should take into consideration the consequences of its RA decisions on the other tiers. The second issue is that centralized RA solutions may not be feasible. Hence, there is a need for decentralized solutions for RA in different network tiers.

In this paper, we formulate the RA problem for a two-tier orthogonal frequency division multiple access (OFDMA) wireless network composed of a macrocell overlaid by small cells. The objective of the macrocell is to allocate resources to its macro UEs (MUEs) to satisfy their data rate requirements. In addition, knowing about the existence of small cells, the macrocell allocates the radio resources (i.e., sub-channel and power) to its MUEs in a way that can sustain the highest interference level from the small cells. For this reason, we formulate an optimization problem for the macrocell with an objective that is different from those in the traditional RA problems. Now, since small cells create dead zones around them in the downlink direction, the MUEs should be protected against transmission from the small cells [3], [5]. Hence, knowing about the maximum allowable interference levels for MUEs, the small cells perform RA by solving an optimization problem whose objective function combines both the admission control (AC) and the consumed bandwidth (i.e., number of allocated sub-channels). The objective of the small cell tier is to admit as many small cell UEs (SUEs) as possible at their target data rates and consume the minimum amount of bandwidth. Again, this follows the same notion of tier-awareness by leaving as much bandwidth as possible for other network tiers (e.g., for device-to-device (D2D) communication). For this, an optimization problem is formulated for the small cell tier with the aforementioned objective, given the QoS requirements of SUEs and the interference constraints for the MUEs. Dual decomposition is used to have a decentralized RA and AC problem by decomposing the optimization problem into sub-problems for each small cell to solve. For this, only local channel gain information is used along with some coordination

A. Abdelnasser (email: nasra@cc.umanitoba.ca) and E. Hossain (email: Ekram.Hossain@umanitoba.ca) are with the Department of Electrical and Computer Engineering at the University of Manitoba, Canada. D. I. Kim (email: dikim@skku.ac.kr) is with the School of Information and Communication Engineering at the Sungkyunkwan University (SKKU), Korea.

A preliminary version of the paper [1] has been submitted to IEEE Globecom'14.

with the Home eNB Gateway (HeNB-GW) [6].

The key contributions of this paper can be summarized as follows:

- We develop a complete framework for tier-aware resource allocation in an OFDMA-based two-tier macrocell-small cell network with new objectives, which are different from the traditional sum-power or sum-rate objectives.
- For the macrocell tier, we formulate a resource allocation problem that is aware of the existence of the small cell tier and show that it is a mixed integer nonlinear program (MINLP).
- We prove that the macrocell can solve another alternate optimization problem that yields the optimal solution for the MINLP with polynomial time complexity.
- We compare the proposed method for the macrocell RA problem to the traditional “minimize the total sum-power” problem and show that the proposed method outperforms the traditional one in terms of the average number of admitted SUEs.
- For the small cell tier, we formulate a joint resource allocation and admission control problem that aims at maximizing the number of admitted SUEs and minimizing their bandwidth consumption to accommodate additional tiers, and show that it is an MINLP.
- We offer an upper bound solution to the MINLP through convex relaxation and propose a solution to the convex relaxation that can be implemented in a distributed fashion using dual decomposition.

The rest of this paper is organized as follows. Section II reviews the related work. Section III presents the system model and assumptions for this work. In Section IV, the optimization problems are formulated for both the macrocell tier and the small cell tier, followed by the use of dual decomposition to have a decentralized operation. Numerical results are discussed in Section V and finally Section VI concludes the work. A summary of the important symbols and notations used in the paper is given in Table I.

II. RELATED WORK

The RA problem in OFDMA-based multi-tier cellular networks has been extensively studied in the literature. The authors in [7] studied the RA problem in a multi-tier cellular network to maximize the sum-throughput subject to simple power budget and sub-channel allocation constraints. However, no QoS constraints were imposed. In [8], the RA problem in a femtocell network was modeled, with interference constraints for MUEs, in order to achieve fairness among femtocells. No QoS constraints, however, were imposed for femtocell users. In [9], the RA problem in a two-tier macrocell-femtocell OFDMA network was modeled as a Stackelberg game, where the macrocell acts as the leader and the femtocells act as the followers. However, no interference constraints for MUEs were considered. Also, no QoS constraints were imposed for femtocells. Reference [10] studied the RA problem in a two-tier network composed of macrocells and femtocells which aimed at maximizing the sum-throughput of femtocells subject to total sum-rate constraint for the macrocell. Nevertheless,

no QoS constraints were imposed for femtocells. The authors in [11] studied the RA problem with QoS and interference constraints in a two-tier cellular network and used clustering as a technique to reduce the overall complexity.

In the above works, either no QoS constraints were imposed or the RA problems with QoS constraints were assumed feasible. In other words, admission control [12], which is a technique to deal with infeasibility when it is not possible to support all UEs with their target QoS requirements, was not studied. The authors in [13] proposed a distributed self-organizing RA scheme for a femtocell only network, with the aim of minimizing the total transmit power subject to QoS constraints. It was shown that minimizing the transmit power (which results in reduced interference) may improve throughput.

Several works in the literature have considered the admission control problem. For cellular cognitive radio networks, [14] studied the problem of admission and power control to admit the maximum number of secondary links and maximize their sum-throughput subject to QoS requirements and interference constraints for primary links. However, power control was done centrally. The authors in [15] considered the problem of admission and power control, where the primary users are guaranteed a premium service rate and the secondary users are admitted (as many as possible) so long as the primary users are not affected. In [16], the authors proposed a joint rate and power allocation scheme with explicit interference protection for primary users and QoS constraints for secondary users, where admission control was performed centrally. However, [14]-[16] only considered single-channel systems. The authors in [17] studied the problem of joint rate and power allocation with admission control in an OFDMA-based cognitive radio network subject to QoS requirements for secondary users and interference constraints for primary users. However, resource allocation and admission control were performed centrally. In addition, channels were randomly allocated to secondary users.

In relay networks, [18] studied the problem of power allocation in amplify and forward wireless relay systems for different objectives, where admission control was employed as a first step preceding power control. However, only one channel was considered. In addition, power and admission control were done centrally. In [19], a joint bandwidth and power allocation for wireless multi-user networks with admission control in relay networks was proposed for different system objectives. Unequal chunks of bandwidths were allocated. However, the resource allocation was performed centrally.

For a two-tier small cell network, [20] studied joint admission and power control. Small cells are admitted into the network so long as the QoS of macrocell users is not compromised. Admission and power control were performed in a distributed fashion. However, only a single channel system was considered. Reference [21] proposed a distributed admission control mechanism for load balancing among sub-carriers with multiple QoS classes. In addition, small cells mitigate co-tier and cross-tier interferences using slot allocation of different traffic streams among different sub-carriers. However, no power allocation was performed.

We notice that none of the quoted works considers the

TABLE I
SUMMARY OF THE IMPORTANT SYMBOLS AND NOTATIONS

Symbol	Description
B	Index of the macrocell
\mathbf{d}	Sub-gradient
d^n	Element of the sub-gradient \mathbf{d}
\mathcal{F}	Set of all SUEs
F	Number of SUEs
f	Index of an SUE
\mathcal{F}_s	Set of all SUEs served by small cell s
$g_{i,j}^n$	Channel gain of the link between UE j served by BS i on sub-channel n
g	Lagrange dual function
g_s	Lagrange dual function for small cell s
I_m^n	Maximum tolerable interference level on sub-channel n allocated to MUE m
I_{max}	Upper limit on the maximum tolerable interference level I_m^n
I_{th}	Equal value for the maximum tolerable interference level I_m^n , $\forall m \in \mathcal{M}, n \in \mathcal{N}$
$I_{th,L}$ and $I_{th,H}$	Lower and upper limits on I_{th} in the bisection method
$I_{th,M}$	Mean of $I_{th,L}$ and $I_{th,H}$
\mathcal{L}	Lagrangian function
\mathcal{M}	Set of all MUEs
M	Number of MUEs
m	Index of an MUE
\mathcal{N}	Set of all available sub-channels
N	Number of sub-channels
n	Index of a sub-channel
$N_{a,c}$	Number of allocated sub-channels
\mathcal{N}_m	Set of all sub-channels allocated to MUE m
N_o	Noise power
$P_{i,j}^n$	Power allocated to the link between UE j served by BS i on sub-channel n
$\bar{P}_{i,j}^n$	Actual power allocated to the link between UE j served by BS i on sub-channel n
$P_{B,max}$	Total macrocell power
$P_{s,max}$	Total small cell power
R_B	Coverage radius of macrocell B
R_m	Data rate requirement of MUE m
R_f	Data rate requirement of SUE f
\mathcal{S}	Set of all small cells
S	Number of small cells
s	Index of a small cell
$y_{s,f}$	Admission control variable for SUE f served by small cell s
α^n	Scale factor for small cells transmission powers on sub-channel n
$\gamma_{B,m}^n$	Received SINR of an MUE m served by macrocell B on sub-channel n
$\gamma_{s,f}^n$	Received SINR of an SUE f served by small cell s on sub-channel n
$\Gamma_{i,j}^n$	Sub-channel allocation indicator for sub-channel n allocated to UE j served by BS i
Δf	Sub-channel bandwidth
δ	Termination tolerance in bisection method
η	Lagrange multiplier associated with the cross-tier interference
η^n	Element of the Lagrange multiplier η
ϵ	Weighting factor

interaction between the different network tiers and the consequences of RA decisions of one tier on the other one. In addition, it is desirable to have an RA and AC scheme that is implementable in a distributed fashion in a dense multi-tier OFDMA network. Table II summarizes the related work and their differences from the work presented in this paper.

III. SYSTEM MODEL, ASSUMPTIONS, AND RESOURCE ALLOCATION FRAMEWORK

A. System Model and Assumptions

We consider the downlink of a two-tier cellular network, where a single macrocell, referred to by the index B and with coverage radius R_B , is overlaid with S small cells. Denote by \mathcal{S} the set of small cells, where $S = |\mathcal{S}|$. A closed-access scheme is assumed for all small cells, where access to a small cell is restricted only to the registered SUEs. All small cells are connected to the mobile core network. For example, femtocells can connect to the core network by using the DSL or CATV modems via an intermediate entity called the Femto Gateway (FGW) or HeNB-GW [6] which can take part in the resource allocation operation for femtocells.

We denote by \mathcal{M} the set of MUEs served by the macrocell B with $M = |\mathcal{M}|$. Each MUE m has a data rate requirement

of R_m . In addition, denote by \mathcal{F} the set of SUEs in the system with $F = |\mathcal{F}|$. Each SUE f has a data rate requirement of R_f . We refer to the set of SUEs served by small cell s by \mathcal{F}_s . We assume that all UEs are already associated with their BSs and that this association remains fixed during the runtime of the resource allocation process. We have $\bigcup_{s=1}^S \mathcal{F}_s = \mathcal{F}$ and $\bigcap_{s=1}^S \mathcal{F}_s = \emptyset$. All MUEs exist outdoor and all SUEs exist indoor. We have an OFDMA system, where we denote by \mathcal{N} the set of available sub-channels with $N = |\mathcal{N}|$ and Δf is the bandwidth of a sub-channel n . Universal frequency reuse is assumed, where the macrocell and all the small cells have access to the set of sub-channels \mathcal{N} . $\Gamma_{i,j}^n$ is the sub-channel allocation indicator, i.e., $\Gamma_{i,j}^n = 1$, if sub-channel n is allocated to UE j served by BS i and takes the value of 0 otherwise.

The UEs are capable of using two modes of sub-channel allocation, namely, the exclusive mode and the time sharing mode. For the exclusive mode, in a given transmission frame, sub-channel n is used by one UE only. In the time sharing mode, a sub-channel n is allocated to a certain UE a portion of the time. In this way, multiple UEs can time share a sub-channel n in a given transmission frame [22].

Denote by $P_{i,j}^n$ and $g_{i,j}^n$ the allocated power to and the

TABLE II
SUMMARY OF RELATED WORK

Previous works	Network type	Objective function	Multi-channel	AC	QoS constraints	Distributed solution	Impact of one tier on another
[7]	Two-tier macrocell/small cell network	Maximize sum-rate for two tiers	Yes	No	No	Yes	No
[8]	Two-tier macrocell/small cell network	Maximize sum-min rate for small cells	Yes	No	For MUEs	Yes	No
[9]	Two-tier macrocell/small cell network	Maximize sum-rate for two tiers	Yes	No	For MUEs	Yes	No
[10]	Two-tier macrocell/small cell network	Maximize sum-rate for small cells	Yes	No	For MUEs	Yes	No
[11]	Two-tier macrocell/small cell network	Maximize sum-rate for small cells	Yes	No	For SUEs	Semi-distributed	No
[13]	Single-tier small cell network	Minimize sum-power	Yes	No	Yes	Yes	No
[14]	Cognitive radio networks	Maximize number of links with max sum-rate	No	Yes	Yes	No	No
[15]	Cognitive radio networks	Maximize number of users with min sum-power	No	Yes	Yes	Yes	No
[16]	Cognitive radio networks	Maximize min-rate and maximize sum-log rate	No	Yes	Yes	No	No
[17]	Cognitive radio networks	Maximize sum-rate	Yes	Yes	Yes	No	No
[18]	Relay networks	Maximize min-SINR, min max-power and max sum-rate	No	Yes	Yes	No	No
[19]	Relay networks	Maximize sum-rate, max min-rate and minimize sum-power	Yes	Yes	Yes	No	No
[20]	Two-tier macrocell/small cell network	Minimize sum-power with max number of SUEs	No	Yes	For MUEs and SUEs	Yes	No
[21]	Two-tier macrocell/small cell network	Maximize product of minimum of ($2 \times$ target rate - achieved rate) and achieved rate	Yes	Yes	For SUEs	Yes	No
Our proposed scheme	Two-tier macrocell/small cell network	Maximize sum-tolerable interference for MUEs and maximize admitted SUEs with minimum bandwidth	Yes	Yes	For MUEs and SUEs	Yes	Yes

channel gain of the link between BS i and UE j on sub-channel n . Channel gains are time varying and account for path-loss, log-normal shadowing, and fast fading. The channel gains are assumed to remain static during the resource allocation process. The received signal to interference plus noise ratio (SINR) $\gamma_{B,m}^n$ of an MUE m served by macrocell B on a sub-channel n is defined as:

$$\gamma_{B,m}^n = \frac{P_{B,m}^n g_{B,m}^n}{I_m^n + N_o} \quad (1)$$

where I_m^n is the maximum tolerable interference level at MUE m on sub-channel n and N_o is the noise power. According to (1), the following constraint holds for small cell transmission powers on sub-channel n :

$$\Gamma_{B,m}^n \left(\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \Gamma_{s,f}^n P_{s,f}^n g_{s,m}^n \right) \leq \Gamma_{B,m}^n I_m^n \quad (2)$$

where the constraint is active only if sub-channel n is allocated to MUE m , i.e., $\Gamma_{B,m}^n = 1$. Similarly, we can define the received SINR $\gamma_{s,f}^n$ of an SUE f served by small cell s on a sub-channel n as:

$$\gamma_{s,f}^n = \frac{P_{s,f}^n g_{s,f}^n}{\sum_{m=1}^M \Gamma_{B,m}^n P_{B,m}^n g_{B,f}^n + N_o}. \quad (3)$$

In (3), we consider cross-tier interference from macrocell B .¹ On the other hand, co-tier interference from other small cells is assumed to be a part of the noise power N_o due to the wall penetration loss and their relatively low transmission powers [23].

B. Tier-Aware Resource Allocation Framework

Fig. 1 describes the RA framework proposed in this paper. Given the rate requirements for the MUEs, the macrocell starts by allocating resources to its MUEs and specifies the maximum tolerable interference levels on each allocated sub-channel. The macrocell then sends this RA information to the HeNB-GW which broadcasts it to the small cells. The small cells then perform RA and AC for its SUEs. For the resulting resource allocation for small cells, the MUEs perform interference measurements and report them to the macrocell BS.

¹Note that it is straightforward to account for interference from other macrocells in (1) and (3). Nevertheless, since we are focusing on the interaction between RA decisions of the macrocell tier represented by macrocell B and the overlaying small cells tier, interference from other macrocells will appear as a constant term in (1) and (3). Besides, I_m^n will represent the maximum tolerable interference from the small cell tier.

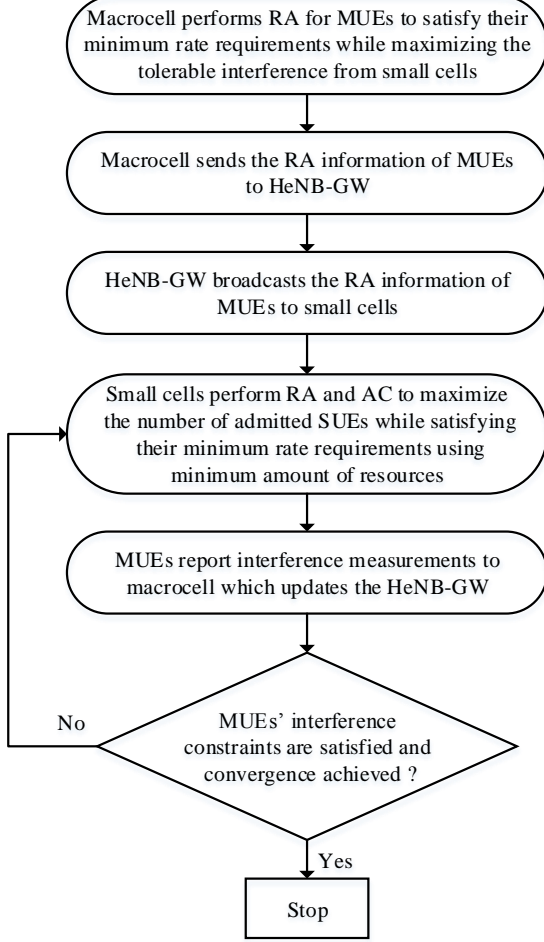


Fig. 1. The RA framework for the macrocell and the small cells.

The macrocell BS then updates the HeNB-GW and the cycle repeats until the interference thresholds for all the MUEs are not violated and the RA and AC converge for all small cells. This cycle repeats due to the distributed nature of RA and AC in small cells. This repetition, however, does not take place if RA and AC in small cells are performed by a central controller. Note that the resource allocation in the macrocell from the first step remains fixed throughout the entire operation of the resource allocation process in the macrocell and small cells. The awareness of the macrocell about the small cell tier is reflected in the way the radio resources are allocated in the macrocell. The macrocell allocates resources to its MUEs in a way that can tolerate the maximum interference possible from the small cell tier. Note, however, that the minimum rate constraints of all MUEs must be satisfied in the sense that the rate requirement for none of the MUEs is compromised for admitting new SUEs. On the other hand, the awareness of the small cell tier about the existence of other tiers is reflected in the fact that the resource allocation in the small cell tier satisfies the rate requirements of the SUEs using the minimum amount of bandwidth resources.

IV. PROBLEM FORMULATIONS FOR RESOURCE ALLOCATION

A. Problem Formulation for Macrocell

The macrocell is responsible for providing the basic coverage for the MUEs [24]. Hence, the target of the macrocell is to allocate the resources to its MUEs to satisfy their data rate requirements and specify the maximum tolerable interference level by its MUEs on the allocated sub-channels. Different methods (i.e., corresponding to optimization problems for resource allocation with different objective functions), as will be shown later, can be followed to accomplish this task. It is of interest to study and understand the effects of different RA methods on other network tiers, which will be the small cell tier in our case.

1) Maximize the sum of tolerable interference levels:

One way of performing RA in the macrocell is to allocate resources to the MUEs in a way that maximizes the sum of the maximum tolerable interference levels on the allocated sub-channels. The motivation behind this objective is to allow the maximum possible freedom for the small cell tier in using the sub-channels. In this context, equal transmit power is assumed on the allocated sub-channels in the macrocell, i.e., $P_{B,m}^n = \frac{P_{B,max}}{N_{ac}}$ [25], where $P_{B,max}$ is the total macrocell power and $N_{ac} \leq N$ is the number of allocated sub-channels. Denote by \mathcal{N}_m the set of sub-channels allocated to MUE m . Hence, we can define the optimization problem in (4), where the objective is to maximize the sum of the tolerable interference levels I_m^n for all MUEs m on all sub-channels n . C1 is the data rate constraint for each MUE m . C2 and C3 indicate that the sets of sub-channels allocated to the MUEs are disjoint and constitute the entire set of sub-channels \mathcal{N} . C4 is a constraint added for numerical purposes, where I_{max} is a very large number and $I_m^n = I_{max}$ means that sub-channel n is not actually allocated to MUE m . Finally, C5 indicates that I_m^n should be positive.

$$\begin{aligned}
 & \max_{\{I_m^n\}} \sum_{m=1}^M \sum_{n \in \mathcal{N}_m} I_m^n \\
 & \text{subject to} \\
 & \text{C1 : } \sum_{n \in \mathcal{N}_m} \Delta f \log_2 \left(1 + \frac{P_{B,m}^n g_{B,m}^n}{I_m^n + N_o} \right) \geq R_m, \forall m \in \mathcal{M} \\
 & \text{C2 : } \mathcal{N}_i \cap \mathcal{N}_j = \emptyset, \forall i, j \in \mathcal{M} \\
 & \text{C3 : } \bigcup_{m=1}^M \mathcal{N}_m \subseteq \{1, 2, \dots, N\} \\
 & \text{C4 : } I_m^n \leq I_{max}, \forall m \in \mathcal{M}, n \in \mathcal{N} \\
 & \text{C5 : } I_m^n \geq 0, \forall m \in \mathcal{M}, n \in \mathcal{N}.
 \end{aligned} \tag{4}$$

In general, (4) is an MINLP whose feasible set is non-convex due to C1 and the combinatorial nature of sub-channel allocation. Besides, $P_{B,m}^n$ is unknown as the number of allocated sub-channels N_{ac} is not known yet. However, by carefully inspecting (4), some interesting features can be revealed which lead to the possibility of obtaining the optimal

solution of (4) with polynomial time complexity. We shall assume first that (4) is always feasible and that in the extreme case, an MUE can have its rate requirement satisfied with one sub-channel only. The last assumption is possible thanks to the fact that the macrocell in our model has a control on the maximum interference level on the allocated sub-channel. The following Lemmas reveal some of the interesting features of (4).

Lemma 1. *At optimality, all data rate constraints C1 hold with equality.*

Proof: Since the objective function in (4) is monotonically increasing in I_m^n and C1 is monotonically decreasing in I_m^n , C1 must hold with equality at optimality for all MUEs. ■

Lemma 2. *At optimality, each MUE m is assigned a single sub-channel i with $I_m^i < I_{max}$.*

Proof: To establish this result, we assume that for an MUE m at optimality, $I_m^n < I_{max}$, $\forall n \in \mathcal{N}_m$ with an objective function value $\text{Obj}_m = \sum_{n \in \mathcal{N}_m} I_m^n$ for MUE m . However, according to Lemma 1, the objective function is monotonically increasing in I_m^n , whereas the constraint C1 is monotonically decreasing in I_m^n . Therefore, we can decrease the value of I_m^i on a certain sub-channel $i \in \mathcal{N}_m$ and increase the values of all other I_m^j , $j \in \mathcal{N}_m, j \neq i$. In this way, we end up with $I_m^j = I_{max}$, for $j \in \mathcal{N}_m, j \neq i$. Meanwhile, I_m^i reaches a value I_m^{i*} such that the rate constraint for MUE m is met with equality resulting in a new objective function value $\text{Obj}_m' = ((|\mathcal{N}_m| - 1)I_{max} + I_m^{i*})$ which is clearly higher than Obj_m . Hence, the initial assumption of optimality is contradicted. ■

Recall that, a value of I_{max} for a certain I_m^n means that sub-channel n is not actually allocated to MUE m . This leads to the fact that a system with N sub-channels and M MUEs, where $M \leq N$, will end up with M sub-channels only allocated to the M MUEs which leads to a minimal use of the available system bandwidth. Hence, $P_{B,m}^n = \frac{P_{B,max}}{N_{ac}} = \frac{P_{B,max}}{M}$.

Lemma 3. *The allocated sub-channel i for MUE m is the one with the highest channel gain $g_{B,m}^i$, $i \in \mathcal{N}_m$.*

Proof: According to Lemma 2, at optimality, $\text{Obj}_m' = ((|\mathcal{N}_m| - 1)I_{max} + I_m^{i*})$ for MUE m , where I_m^{i*} is selected such that the achieved data rate on sub-channel i is equal to R_m . Hence, from the rate constraint formula, $I_m^{i*} = \left(\frac{P_{B,m}^i g_{B,m}^i}{2^{R_m/\Delta f} - 1} - N_o \right)$. It is clear that I_m^{i*} is directly proportional to $g_{B,m}^i$. Therefore, to maximize Obj_m' , we need to maximize I_m^{i*} . Hence, MUE m should be allocated sub-channel i such that $i = \arg \max_{n \in \mathcal{N}_m} g_{B,m}^n$. ■

Based on the given Lemmas, we have the following Theorem.

Theorem 1. *To maximize the sum of tolerable interference levels, the macrocell can solve the following alternate opti-*

mization problem:

$$\begin{aligned} & \max_{\{\Gamma_{B,m}^n\}} \sum_{m=1}^M \sum_{n=1}^N \Gamma_{B,m}^n g_{B,m}^n \\ & \text{subject to} \\ & \text{C1 : } \sum_{n=1}^N \Gamma_{B,m}^n = 1, \forall m \in \mathcal{M} \\ & \text{C2 : } \sum_{m=1}^M \Gamma_{B,m}^n \leq 1, \forall n \in \mathcal{N} \\ & \text{C3 : } \Gamma_{B,m}^n \in \{0, 1\}, \forall m \in \mathcal{M}, n \in \mathcal{N} \end{aligned} \quad (5)$$

where the objective in (5) is to maximize the sum of the allocated sub-channel gains. C1 restricts the number of allocated sub-channels to any MUE m to one sub-channel only, whereas C2 restricts sub-channel n to be allocated to at most one MUE. Then for each MUE m with allocated sub-channel n , the maximum tolerable interference level is given by: $I_m^n = \left(\frac{P_{B,m}^n g_{B,m}^n}{2^{R_m/\Delta f} - 1} - N_o \right)$. Hence, (4) is solved optimally.

Proof: According to Lemmas 2 and 3, at optimality, each MUE will have only one sub-channel which is the one with the highest gain. Hence, we can define the optimization problem in (5) which is the well-known assignment problem that can be efficiently solved in polynomial time using the Hungarian method [26]. ■

In this way, the macrocell allocates sub-channels to its MUEs in a way that satisfies their data rate requirements and that can tolerate the maximum possible interference from the small cell tier.

2) *Minimize the total sum-power:* As a baseline, we consider another way of performing RA in the macrocell by minimizing the total sum-power of the macrocell given the data rate requirements of the MUEs. This problem has been studied extensively in the literature [27]. However, the formulation developed in [27] does not account for the maximum tolerable interference level I_m^n . Hence, we include it here with the required modification to determine the maximum tolerable interference level I_m^n . We have, thus, the following optimization problem:

$$\begin{aligned} & \min_{\{P_{B,m}^n\}} \sum_{m=1}^M \sum_{n \in \mathcal{N}_m} P_{B,m}^n \\ & \text{subject to} \\ & \text{C1 : } \sum_{n \in \mathcal{N}_m} \Delta f \log_2 \left(1 + \frac{P_{B,m}^n g_{B,m}^n}{I_m^n + N_o} \right) \geq R_m, \forall m \in \mathcal{M} \\ & \text{C2 : } \mathcal{N}_i \cap \mathcal{N}_j = \emptyset, \forall i, j \in \mathcal{M} \\ & \text{C3 : } \bigcup_{m=1}^M \mathcal{N}_m \subseteq \{1, 2, \dots, N\} \\ & \text{C4 : } P_{B,m}^n \geq 0, \forall m \in \mathcal{M}, n \in \mathcal{N}. \end{aligned} \quad (6)$$

In (6), given the maximum tolerable interference level on each allocated sub-channel I_m^n , the macrocell seeks a power and sub-channel allocation solution that minimizes the sum-power. Although (6) is an MINLP whose feasible set is non-convex, it has been solved efficiently in [27] in the dual domain using dual decomposition relying on the fact that the duality gap becomes virtually zero when the number of sub-channels in the system is sufficiently large.

Remark 1. *The reason for choosing the resource allocation solution based on the formulation in (6) as the baseline is the following. With this solution, at optimality, all MUEs have their rate requirements satisfied with equality. This is also the case for the solution obtained from the formulation in (4). In other words, with both the solutions, the MUEs achieve the same performance. The difference however lies in the way the resources are allocated, which subsequently impacts the performance of the small cell tier.*

In (6), the same value of I_m^n is assumed $\forall m \in \mathcal{M}, n \in \mathcal{N}$, i.e., $I_m^n = I_{th}$. For a further fair comparison between (4) and (6), the macrocell adjusts the maximum tolerable interference level I_{th} such that $\sum_{m=1}^M \sum_{n \in \mathcal{N}_m} P_{B,m}^n = P_{B,max}$. This can be accomplished by using the bisection method according to **Algorithm 1** as given below, where $I_{th,H} > I_{th,L}$.

Algorithm 1 Bisection method to find optimal I_{th}

- 1: Macrocell initializes $I_{th,L}$, $I_{th,H}$, and δ
 - 2: **while** $|\sum_{m=1}^M \sum_{n \in \mathcal{N}_m} P_{B,m}^n - P_{B,max}| > \delta$ **do**
 - 3: $I_{th,M} = (I_{th,L} + I_{th,H}) / 2$
 - 4: Macrocell solves the optimization problem in (6)
 - 5: **if** $\sum_{m=1}^M \sum_{n \in \mathcal{N}_m} P_{B,m}^n > P_{B,max}$ **then**
 - 6: $I_{th,H} = I_{th,M}$
 - 7: **else if** $\sum_{m=1}^M \sum_{n \in \mathcal{N}_m} P_{B,m}^n < P_{B,max}$ **then**
 - 8: $I_{th,L} = I_{th,M}$
 - 9: **end if**
 - 10: **end while**
-

After **Algorithm 1** terminates, $I_{th,M}$ gives the optimal value of I_{th} . The optimization problem in (6) ends up with the power and sub-channel allocation to the MUEs with a uniform maximum tolerable interference level I_{th} on all allocated sub-channels. In general, as will be shown in the numerical results, (6) leads to a higher number of allocated sub-channels to the MUEs than that (4) does. It is of interest to study the effect of the two different RA methods on the small cell tier.

B. Problem Formulation for Small Cells

Due to the small distance and the good channel conditions between small cells and SUEs, small cells are capable of serving registered SUEs with higher data rates than the macrocell. However, this should not be at the cost of QoS degradation at MUEs as they are served by the macrocell and provided with basic coverage at possibly lower rates [10]. Hence, given the maximum tolerable interference levels on each allocated sub-channel for the MUEs, each small cell now tries to admit as many SUEs as possible at their target data rate by using the minimum possible bandwidth. Again, the idea here is to

leave as much bandwidth as possible for the other network tiers (e.g., for device-to-device (D2D) communication).

1) *Centralized operation:* To accomplish the aforementioned requirements, we define the optimization problem in (7), where the objective function accounts for both admission control and sub-channel allocation. We have the admission control variable $y_{s,f}$ which takes the value of 1 if SUE f is admitted in small cell s and 0 otherwise. By controlling the weighting factor $\epsilon \in [0, 1]$, admission control can be given higher priority over the number of used sub-channels.

$$\begin{aligned}
 & \max_{\{\Gamma_{s,f}^n, P_{s,f}^n, y_{s,f}\}} (1 - \epsilon) \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} y_{s,f} - \epsilon \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \Gamma_{s,f}^n \\
 & \text{subject to} \\
 & \text{C1: } \sum_{n=1}^N \Delta f \log_2 \left(1 + \frac{P_{s,f}^n g_{s,f}^n}{\sum_{m=1}^M \Gamma_{B,m}^n P_{B,m}^n g_{B,f}^n + N_o} \right) \\
 & \quad \geq y_{s,f} R_f, \quad \forall s \in \mathcal{S}, f \in \mathcal{F}_s \\
 & \text{C2: } \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N P_{s,f}^n \leq P_{s,max}, \quad \forall s \in \mathcal{S} \\
 & \text{C3: } \Gamma_{B,m}^n \left(\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} P_{s,f}^n g_{s,m}^n \right) \leq \Gamma_{B,m}^n I_m^n, \quad \forall n \in \mathcal{N} \\
 & \text{C4: } P_{s,f}^n \leq \Gamma_{s,f}^n P_{s,max} \quad \forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N} \\
 & \text{C5: } \sum_{f \in \mathcal{F}_s} \Gamma_{s,f}^n \leq 1, \quad \forall s \in \mathcal{S}, n \in \mathcal{N} \\
 & \text{C6: } P_{s,f}^n \geq 0, \quad \forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N} \\
 & \text{C7: } \Gamma_{s,f}^n \in \{0, 1\}, \quad \forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N} \\
 & \text{C8: } y_{s,f} \in \{0, 1\}, \quad \forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N}. \tag{7}
 \end{aligned}$$

Proposition 1. *By choosing $\epsilon < \frac{1}{1+SN}$, (7) admits the maximum number of SUEs while consuming the minimum number of sub-channels.*

Proof: This Proposition can be proved in a way similar to that in [28]. Let $(\Gamma_{s,f}^{n*}, P_{s,f}^{n*}, y_{s,f}^*)$, $\forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N}$ denote an optimal solution of (7). Let $(\hat{\Gamma}_{s,f}^n, \hat{P}_{s,f}^n, \hat{y}_{s,f}^n)$, $\forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N}$ be a feasible solution that admits one more SUE than the optimal solution, i.e., $\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \hat{y}_{s,f}^n = \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} y_{s,f}^* + 1$. The objective of the feasible solution can be written as:

$$(1 - \epsilon) \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \hat{y}_{s,f}^n - \epsilon \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \hat{\Gamma}_{s,f}^n \stackrel{(1)}{\geq}$$

$$(1 - \epsilon) \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} y_{s,f}^* + (1 - \epsilon) - \epsilon SN \stackrel{(2)}{\geq}$$

$$(1 - \epsilon) \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} y_{s,f}^* \stackrel{(3)}{\geq}$$

$$(1 - \epsilon) \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} y_{s,f}^* - \epsilon \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \Gamma_{s,f}^{n*}.$$

The first inequality holds due to the fact that $\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \Gamma_{s,f}^n$ is upper bounded by SN when all sub-channels in all small cells are allocated. The second inequality holds by setting $(1 - \epsilon) - \epsilon SN > 0$. Hence, we have $\epsilon < \frac{1}{1+SN}$. The last inequality holds due to the non-negativity of $\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \Gamma_{s,f}^{n*}$. In this way, the value of the objective function for the feasible solution is higher than the optimal one, which contradicts the optimality of $(\Gamma_{s,f}^{n*}, P_{s,f}^{n*}, y_{s,f}^*)$. Thus, there is no other solution that admits a higher number of SUEs under the constraint in (7). Given the optimum value for the admission control variable $y_{s,f}^*$, (7) reduces to a feasible sub-channel and power allocation problem with respect to the variables $\Gamma_{s,f}^n$ and $P_{s,f}^n$ that aims at minimizing the number of used sub-channels subject to the given constraints. ■

In (7), C1 is a data rate constraint for an SUE f which is active only if SUE f is admitted, i.e., $y_{s,f} = 1$. C2 is the power budget constraint for each small cell s restricting the total transmission power of small cell s to be less than or equal to $P_{s,max}$. C3 is a constraint on the maximum cross-tier interference introduced to MUE m using sub-channel n . C4 ensures that if sub-channel n is not allocated to SUE f , its corresponding transmit power $P_{s,f}^n = 0$. C5 constrains sub-channel n to be allocated to at most one SUE f in small cell s . C6 ensures that the power $P_{s,f}^n$ should be positive, and finally, C7 and C8 indicate that $\Gamma_{s,f}^n$ and $y_{s,f}$ are binary variables.

Claim 1. *The optimization problem in (7) is always feasible.*

Proof: A trivial feasible solution of (7) is $\Gamma_{s,f}^n = 0, P_{s,f}^n = 0$ and $y_{s,f} = 0, \forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N}$. ■

The problem in (7) is a MINLP whose feasible set is non-convex due to the combinatorial nature of sub-channel allocation and admission control. However, for small-sized problems, we use OPTI [29], which is a MATLAB toolbox to construct and solve linear, nonlinear, continuous and discrete optimization problems, to obtain the optimal solution. Obtaining the optimal solution, however, for larger problems is intractable. Another approach that can render the problem in (7) more tractable is to have a convex reformulation of (7) by relaxing the constraints C7 and C8 and allowing $\Gamma_{s,f}^n$ and $y_{s,f}$ to take any value in the range $[0, 1]$. Thus, $\Gamma_{s,f}^n$ is now a time sharing factor that indicates the portion of time sub-channel n is allocated to SUE f [30], [31], whereas $y_{s,f}$ indicates the ratio of the achieved data rate for SUE f . Hence, we define the convex optimization problem in (8), where $\tilde{P}_{s,f}^n$ can be related to $P_{s,f}^n$ in (7) as $\tilde{P}_{s,f}^n = \Gamma_{s,f}^n P_{s,f}^n$ to denote the actual transmit power [22]. Now, (8) is a convex optimization problem with a linear objective function and convex feasible set. It can be solved efficiently by the interior point method [32].

Note that the system model assumed in (8) differs from the original one in (7) as it allows time sharing of sub-channels and fractional satisfaction of the required data rates. Hence, the solution of (8) gives an upper bound to the optimal solution of (7). However, it helps by revealing some insights about the

behavior of (7). Note that the solution of (8) necessitates the existence of a central controller which can be, for example, the HeNB-GW. However, for dense small cell networks, having a decentralized solution with some coordination with a central entity will be a more viable option.

$$\begin{aligned} & \max_{\{\Gamma_{s,f}^n, \tilde{P}_{s,f}^n, y_{s,f}\}} (1 - \epsilon) \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} y_{s,f} - \epsilon \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \Gamma_{s,f}^n \\ & \text{subject to} \\ & \text{C1: } \sum_{n=1}^N \Gamma_{s,f}^n \Delta f \log_2 \left(1 + \frac{(\tilde{P}_{s,f}^n g_{s,f}^n / \Gamma_{s,f}^n)}{\sum_{m=1}^M \Gamma_{B,m}^n P_{B,m}^n g_{B,f}^n + N_o} \right) \\ & \quad \geq y_{s,f} R_f, \quad \forall s \in \mathcal{S}, f \in \mathcal{F}_s \\ & \text{C2: } \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \tilde{P}_{s,f}^n \leq P_{s,max}, \quad \forall s \in \mathcal{S} \\ & \text{C3: } \Gamma_{B,m}^n \left(\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \tilde{P}_{s,f}^n g_{s,m}^n \right) \leq \Gamma_{B,m}^n I_m^n, \quad \forall n \in \mathcal{N} \\ & \text{C4: } \sum_{f \in \mathcal{F}_s} \Gamma_{s,f}^n \leq 1, \quad \forall s \in \mathcal{S}, n \in \mathcal{N} \\ & \text{C5: } \tilde{P}_{s,f}^n \geq 0, \quad \forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N} \\ & \text{C6: } \Gamma_{s,f}^n \in (0, 1], \quad \forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N} \\ & \text{C7: } y_{s,f} \in [0, 1], \quad \forall s \in \mathcal{S}, f \in \mathcal{F}_s, n \in \mathcal{N}. \end{aligned} \quad (8)$$

2) *Distributed operation:* To fulfill the requirement of having a decentralized solution for (8), we use the dual decomposition method [33]. For this purpose, we define the following partial Lagrangian function of the primal problem in (8) formed by dualizing the constraint C3:

$$\begin{aligned} & \mathcal{L}(\Gamma_{s,f}^n, \tilde{P}_{s,f}^n, y_{s,f}, \boldsymbol{\eta}) \\ & = (1 - \epsilon) \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} y_{s,f} - \epsilon \sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \Gamma_{s,f}^n \\ & \quad + \sum_{n=1}^N \eta^n \left(\Gamma_{B,m}^n I_m^n - \Gamma_{B,m}^n \left(\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \tilde{P}_{s,f}^n g_{s,m}^n \right) \right) \end{aligned} \quad (9)$$

where $\boldsymbol{\eta}$ is the Lagrange multiplier vector (with elements η^n) associated with the cross-tier interference constraint C3. Then the Lagrange dual function is represented as

$$\begin{aligned} & g(\boldsymbol{\eta}) = \max_{\{\Gamma_{s,f}^n, \tilde{P}_{s,f}^n, y_{s,f}\}} \mathcal{L}(\Gamma_{s,f}^n, \tilde{P}_{s,f}^n, y_{s,f}, \boldsymbol{\eta}) \\ & \text{subject to} \\ & \text{C1, C2, C4 - C7.} \end{aligned} \quad (10)$$

From (9), the maximization of \mathcal{L} can be decomposed into S independent optimization problems for each small cell s as

follows:

$$g_s(\boldsymbol{\eta}) = \max_{\{\Gamma_{s,f}^n, \tilde{P}_{s,f}^n, y_{s,f}\}} (1 - \epsilon) \sum_{f \in \mathcal{F}_s} y_{s,f} - \epsilon \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \Gamma_{s,f}^n - \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \eta^n \Gamma_{B,m}^n \tilde{P}_{s,f}^n g_{s,m}^n$$

subject to

$$C1, C2, C4 - C7, \quad \forall s \in S. \quad (11)$$

Thus, the Lagrange dual function is

$$g(\boldsymbol{\eta}) = \sum_{s=1}^S g_s(\boldsymbol{\eta}) + \sum_{n=1}^N \eta^n \Gamma_{B,m}^n I_m^n. \quad (12)$$

Then, the dual problem is given by:

$$\min_{\eta \geq 0} g(\boldsymbol{\eta}). \quad (13)$$

In order to solve the dual problem, $\boldsymbol{\eta}$ can be updated efficiently using the ellipsoid method [34]. A sub-gradient \mathbf{d} of this problem required for the ellipsoid method is derived in the following proposition.

Proposition 2. *For the optimization problem in (8) with a dual objective $g(\boldsymbol{\eta})$ defined in (10), the following choice of \mathbf{d}^n is a sub-gradient for $g(\boldsymbol{\eta})$:*

$$\mathbf{d}^n = \Gamma_{B,m}^n I_m^n - \Gamma_{B,m}^n \left(\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \tilde{P}_{s,f}^{n*} g_{s,m}^n \right) \quad (14)$$

where \mathbf{d}^n is an element of \mathbf{d} and $\Gamma_{s,f}^{n*}$, $\tilde{P}_{s,f}^{n*}$, and $y_{s,f}^*$ optimize the maximization problem in the definition of $g(\boldsymbol{\eta})$.

Proof: For any $\boldsymbol{\xi} \geq 0$,

$$\begin{aligned} g(\boldsymbol{\xi}) &\geq \mathcal{L}(\Gamma_{s,f}^{n*}, \tilde{P}_{s,f}^{n*}, y_{s,f}^*, \boldsymbol{\xi}) \\ &= g(\boldsymbol{\eta}) + \sum_{n=1}^N (\xi^n - \eta^n) \left[\Gamma_{B,m}^n I_m^n - \Gamma_{B,m}^n \left(\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \tilde{P}_{s,f}^{n*} g_{s,m}^n \right) \right]. \end{aligned}$$

Algorithm 2 gives a practical implementation of the distributed joint RA and AC operation for the small cells. After the macrocell has performed RA for its MUEs, it sends the sub-channel allocation information for its MUEs and the initialized multiplier $\boldsymbol{\eta}$ to the HeNB-GW. For a given $\boldsymbol{\eta}$, all small cells solve their optimization problem in (11) simultaneously. For the given resource allocation in the small cells, the MUEs estimate the resulting interference levels and send them to the macrocell which updates the multiplier values using the ellipsoid method. The macrocell then informs the updated multiplier values to the HeNB-GW, which broadcasts them to the small cells, and the entire operation repeats. Note that the small cells can obtain the channel gains $g_{s,m}^n$ relying on the techniques proposed in [7].

Finally, the remaining issue is to obtain a feasible primal solution to (8) based on the resulting solution from the

Lagrangian dual in (13). It has been reported in [33] and [35] that the iterations of the dual decomposition method are, in general, infeasible with respect to (8). This infeasibility, however, is not severe as large constraint violations usually get penalized. Hence, using a simple procedure, one can recover a primal feasible solution that serves as a lower bound for the optimal solution of (8). Suppose that the reported interference level by an MUE m allocated a sub-channel n was found to be:

$$\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \tilde{P}_{s,f}^n g_{s,m}^n = \alpha^n I_m^n, \quad \alpha^n > 1. \quad (15)$$

A straightforward way to recover feasibility is for the HeNB-GW to instruct all small cells transmitting on sub-channel n to scale down their transmission powers by the factor α^n . For the updated power values, the entire problem is solved to obtain the updated values of sub-channel allocation and admission control variables. The gap between the lower bound offered by this procedure and the upper bound offered by (13), referred to as the duality gap, diminishes with iterations. Convergence to the optimal solution is guaranteed since the primal optimization problem in (8) is convex.

Algorithm 2 Distributed joint RA and AC algorithm

- 1: Macrocell initializes $\boldsymbol{\eta}$, L_{max} , sends sub-channel allocation information $\Gamma_{B,m}^n$ and $\boldsymbol{\eta}$ to HeNB-GW and sets iteration counter $l = 1$
 - 2: HeNB-GW broadcasts $\Gamma_{B,m}^n$ and $\boldsymbol{\eta}$ values to all small cells
 - 3: **repeat**
 - 4: **for** $s = 1 : S$ **do**
 - 5: All small cells solve (11) in parallel
 - 6: **end for**
 - 7: All MUEs estimate interference levels on allocated sub-channels and report them to the macrocell
 - 8: Macrocell evaluates the sub-gradient (14) and updates $\boldsymbol{\eta}$ using the ellipsoid method
 - 9: Macrocell sends updated $\boldsymbol{\eta}$ to HeNB-GW
 - 10: HeNB-GW broadcasts updated $\boldsymbol{\eta}$ to all small cells
 - 11: Macrocell sets $l = l + 1$
 - 12: **until** Convergence or $l = L_{max}$
-

V. NUMERICAL RESULTS AND DISCUSSIONS

A. Parameters

We evaluate the system performance through extensive simulations under various topologies and scenarios. We have a macrocell located at the origin with radius 300 m. A hot spot of small cells exists at a distance of 100 m from the macrocell. The MUEs exist outdoor in this hot spot and are served by the macrocell. Each small cell has 2 indoor SUEs located randomly on a circular disc around the small cell with an inner radius of 3 m and an outer radius of 10 m [36]. The macrocell has a total power budget of $P_{B,max} = 20$ W.

To model the propagation environment, the channel models from [36] are used. The channel gains include path-loss,

log-normal shadowing, and multipath Rayleigh fading. The path-loss between a small cell and its served SUE, $PL = 38.46 + 20 \log R$ and the path-loss between a small cell and the outdoor MUEs, $PL = \max(38.46 + 20 \log R, 15.3 + 37.6 \log R) + L_{ow}$, where R is the distance between a small cell and the UE and L_{ow} accounts for losses due to walls. For path-loss between the macrocell and an SUE existing indoor, $PL = 15.3 + 37.6 \log R + L_{ow}$ and for path-loss between the macrocell and its MUE, $PL = 15.3 + 37.6 \log R$. We have the following values for the standard deviation of log-normal shadowing: 4 dB for shadowing between SUE and its small cell, 8 dB for shadowing between MUE and small cell and 10 dB for shadowing between macrocell and SUE or MUE. The Rayleigh fading gain is modeled as an exponential random variable with unit mean. We assume $\Delta f = 180$ KHz, $\epsilon = \frac{0.9}{1+SN}$, and noise power, $N_o = 10^{-13}$ W. I_{max} is set to any arbitrary large number. All the rate requirements in the numerical results are specified in terms of spectral efficiency (bps/Hz).

In the numerical results, the following performance metrics are used:

- Average percentage of admitted SUEs $= \frac{\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} y_{s,f}}{F} \times 100$.
- Average percentage of channel usage $= \frac{\sum_{s=1}^S \sum_{f \in \mathcal{F}_s} \sum_{n=1}^N \Gamma_{s,f}^n}{SN} \times 100$.

B. Numerical Results

1) *Comparison between the traditional and the proposed RA method for macrocell:* In this section, we compare the two proposed schemes for RA in the macrocell, namely, the formulation in (4), which we refer to as “proposed” and the formulation in (6), which we refer to as “traditional”. Fig. 2 shows the channel gain realizations for a snapshot of 3 MUEs with 10 sub-channels.

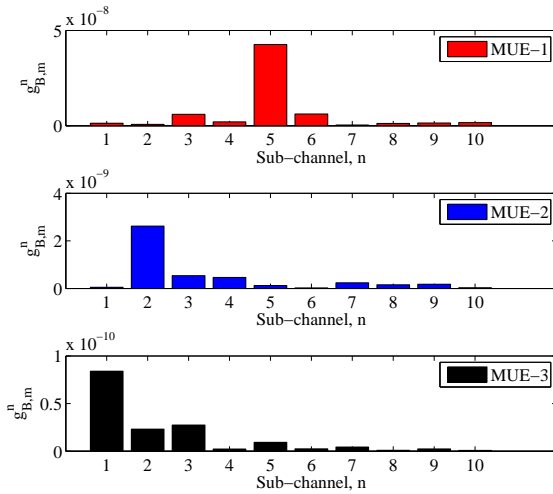


Fig. 2. Channel gains $g_{B,m}^n$ for MUEs $\{1, 2, 3\}$.

Figs. 3-4 compare the two RA results for the given snapshot with $R_m = 5$ bps/Hz. Each figure shows the allocated

power $P_{B,m}^n$ by the macrocell and the maximum tolerable interference level I_m^n on allocated sub-channel n to MUE m . No values for power $P_{B,m}^n$ on the x -axis indicate unallocated sub-channel with the corresponding value for I_m^n set to I_{max} , which means that this sub-channel can be used by the small cell tier unconditionally. For further clarification, Table III shows the absolute values of $P_{B,m}^n$ and I_m^n .

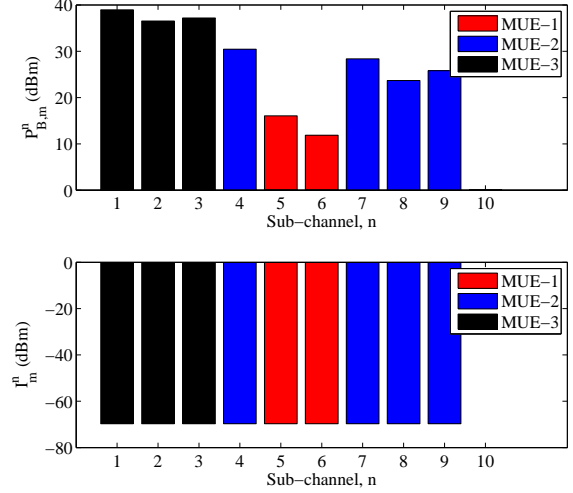


Fig. 3. Allocated power $P_{B,m}^n$ and maximum tolerable interference level I_m^n for MUEs $\{1, 2, 3\}$ using the traditional scheme.

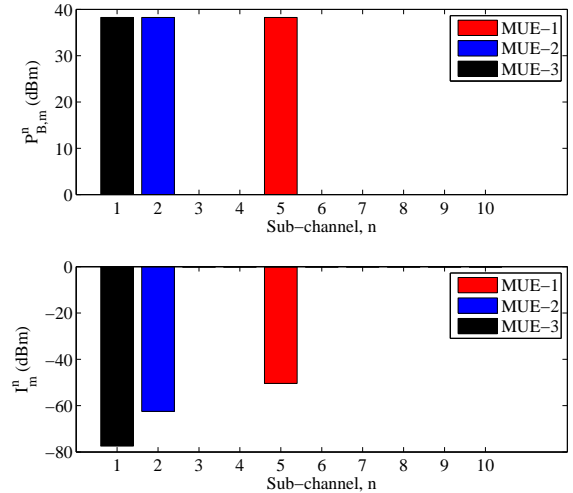


Fig. 4. Allocated power $P_{B,m}^n$ and maximum tolerable interference level I_m^n for MUEs $\{1, 2, 3\}$ using the proposed scheme.

It is clear from Fig. 3 that most of the sub-channels are allocated to the MUEs (9 sub-channels out of 10 are allocated to the 3 MUEs), when using the traditional scheme for RA. We notice also that the macrocell favors good sub-channels as they require less transmit power to achieve the rate requirements for the MUEs, leading at the end to minimum transmit power requirements.

Fig. 4, on the other hand, shows that the 3 MUEs require only 3 sub-channels to achieve their rate requirements, as was proved before, using the proposed scheme. Again, the macrocell allocates the best sub-channels to the MUEs. From Figs. 3-4 and Table III, we can notice that the entire power budget of macro BS, $P_{B,max}$, is used in both cases. It is worth mentioning that when we use the traditional scheme for macrocell resource allocation, it does not necessarily mean that it will consume less power than the proposed scheme, since the maximum tolerable interference level I_m^n is adjusted according to **Algorithm 1** by the macrocell to use the entire power budget. It rather means that, given the maximum tolerable interference levels, the resulting sub-channel and power allocation for the traditional scheme will consume the minimum power and any other allocation will consume a higher power.

Now, for the maximum tolerable interference levels I_m^n , it is obvious from Figs. 3-4 and Table III that the proposed scheme can sustain higher interference levels from the small cell tier. For the traditional scheme, the sum of the tolerable interference is given by: $(9 \times 1.0709 \times 10^{-10}) + I_{max}$. On the other hand, the sum of the tolerable interference levels for the proposed scheme can be given by: $(0.1796 \times 10^{-10}) + (5.6298 \times 10^{-10}) + (91.604 \times 10^{-10}) + (7 \times I_{max})$. It is of interest to compare the effect of the two different RA schemes for the macrocell on the small cell tier.

Fig. 5 compares the average percentage of admitted SUEs when the macrocell performs RA in order to minimize the sum-power (as in (6)) and maximize the sum of tolerable interference levels (as in (4)) with two different wall loss scenarios. We have the following scenario: 2 small cells located at $(-10, -100)$, $(10, -100)$ in a square area hot spot of dimensions 20×20 m², 10 sub-channels, $P_{s,max} = 30$ mW, $R_f = 50$ bps/Hz, and $R_m = 5$ bps/Hz. Numerical results are obtained and averaged for 50 different realizations, where in each realization, the UE positions and the channel gains are varied. The small cell problem is solved centrally using the convex formulation in (8). It is clear from the figure that the proposed RA method for the macrocell outperforms the traditional one. When the macrocell performs RA according to the proposed method, it consumes the minimum bandwidth, and therefore, frees as many sub-channels as possible for the small cells. On the other hand, the traditional method consumes more bandwidth than the proposed one, hence, the small cells have more interference constraints to abide by. We also notice that as the wall losses increase, the small cells tend to be more isolated and the impact of resource allocation in the macrocell on the small cell performance is low.

2) *Comparison between the different formulations for the RA problem for small cells:* Fig. 6 compares the values of the objective function for the MINLP formulation in (7), the centralized convex formulation in (8), and the distributed formulation in (9) for a snapshot of the following scenario: 2 small cells located at $(-10, -100)$, $(10, -100)$ in a square area hot spot of dimensions 20×20 m², 3 sub-channels, 3 MUEs, $P_{s,max} = 30$ mW, $L_{ow} = 1$ dB, and $R_f = 5$ bps/Hz. As was stated previously, the convex formulation provides an upper bound for the solution of the MINLP formulation. Also, we notice that the centralized and distributed formulations have

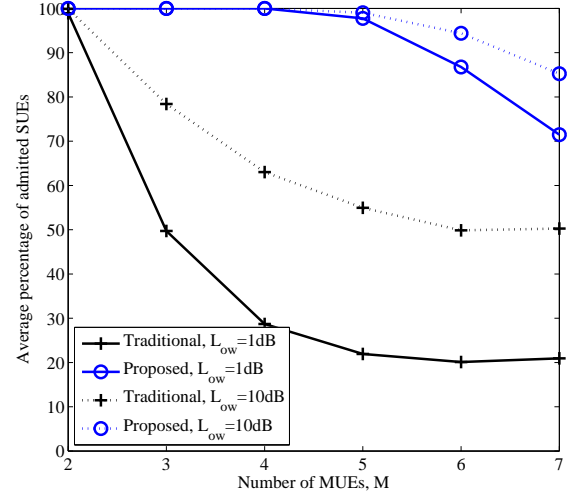


Fig. 5. Average percentage of admitted SUEs vs. number of MUEs M when the macrocell employs both the proposed and the traditional methods for RA with different wall loss scenarios.

the same solution due to the convexity of the centralized formulation in (8). It is worth mentioning that the convex formulation exhibits a behavior similar to the MINLP formulation. Hence, solving the convex formulations reveals insights into the behavior of the solution of the MINLP formulation. We also notice that as R_m increases, the interference constraints for the MUEs become tighter. Hence, the average number of admitted SUEs decreases. Since the objective function in our formulation gives more priority to admission control, the value of objective function decreases with increasing R_m .

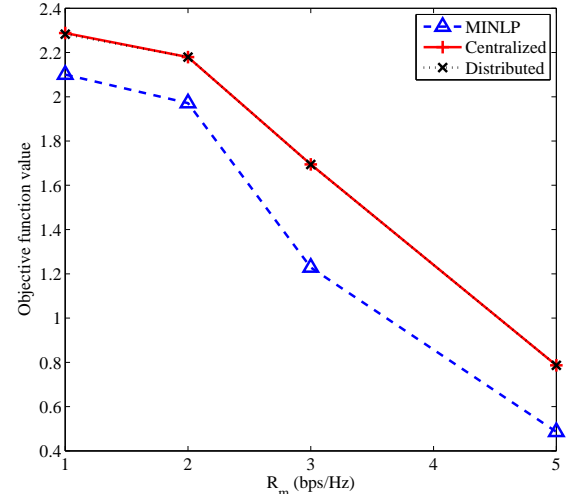


Fig. 6. The values of objective function for different formulations vs. R_m .

Figs. 7-8 show the average percentage of admitted SUEs and channel usage in a small cell vs. R_m for the same scenario considered in Fig. 6. As was discussed in Fig. 6, as the rate requirements for the MUEs increase, they have tighter

TABLE III
ABSOLUTE VALUES OF $P_{B,m}^n$ AND I_m^n FOR THE TRADITIONAL AND PROPOSED MACROCELL RA SCHEMES

			Sub-channel#									
			1	2	3	4	5	6	7	8	9	10
Traditional scheme	$P_{B,m}^n$ (W)	MUE-1	0	0	0	0	0.0402	0.0153	0	0	0	0
		MUE-2	0	0	0	1.1097	0	0	0.6855	0.2333	0.3805	0
		MUE-3	7.8381	4.4890	5.2142	0	0	0	0	0	0	0
	I_m^n (W) $\times 10^{-10}$	MUE-1	I_{max}	I_{max}	I_{max}	I_{max}	1.0709	1.0709	I_{max}	I_{max}	I_{max}	I_{max}
		MUE-2	I_{max}	I_{max}	I_{max}	1.0709	I_{max}	I_{max}	1.0709	1.0709	1.0709	I_{max}
		MUE-3	1.0709	1.0709	1.0709	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}
Proposed scheme	$P_{B,m}^n$ (W)	MUE-1	0	0	0	0	6.6667	0	0	0	0	0
		MUE-2	0	6.6667	0	0	0	0	0	0	0	0
		MUE-3	6.6667	0	0	0	0	0	0	0	0	0
	I_m^n (W) $\times 10^{-10}$	MUE-1	I_{max}	I_{max}	I_{max}	I_{max}	91.604	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}
		MUE-2	I_{max}	5.6298	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}
		MUE-3	0.1796	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}	I_{max}

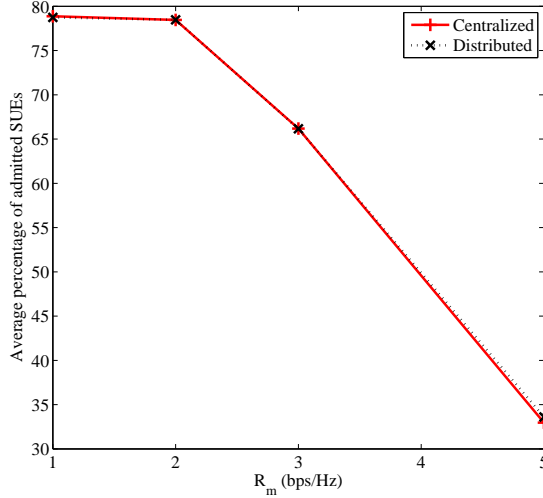


Fig. 7. Average percentage of admitted SUEs vs. R_m .

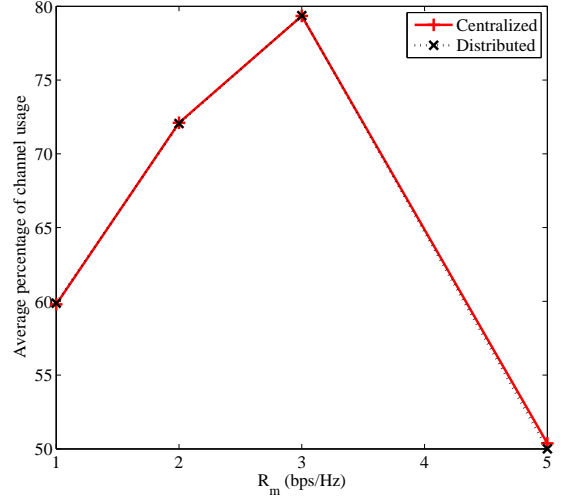


Fig. 8. Average percentage of channel usage vs. R_m .

interference constraints. Hence, the percentage of admitted SUEs generally decreases. We notice in Fig. 7 that, initially, the average percentage of admitted SUEs is almost constant due to the increased number of used sub-channels as shown in Fig. 8. As the MUEs' rate requirements increase further, the increase in the number of used sub-channels is not enough to accommodate the rate requirements of the SUEs, hence, the average percentage of admitted SUEs decreases.

3) *Convergence behavior*: Using the same scenario described for the previous figure, Fig. 9 shows the convergence behavior of **Algorithm 2**, where the upper bound refers to (13) and the lower bound refers to the feasible objective obtained by the procedure at the end of Section IV-B2. In the figure, the best lower bound is obtained by keeping track of the best primal feasible objective resulting through iterations. It is clear that **Algorithm 2** converges to the optimal solution of (8) within a few iterations.

4) *Average percentage of admitted SUEs vs. R_f* : In this scenario, we have the following setup: 5 small cells located at $(-20, -100)$, $(-20, -140)$, $(20, -140)$, $(20, -100)$, $(0, -120)$ in a square area hot spot of dimensions 40×40 m², 5 sub-channels, 5 MUEs, $P_{s,max} = 30$ mW, $L_{ow} = 1$ dB and $R_m = 4$ bps/Hz. Numerical results are obtained and averaged

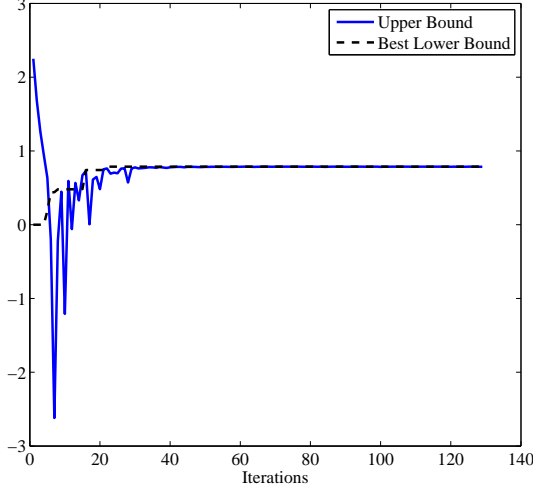
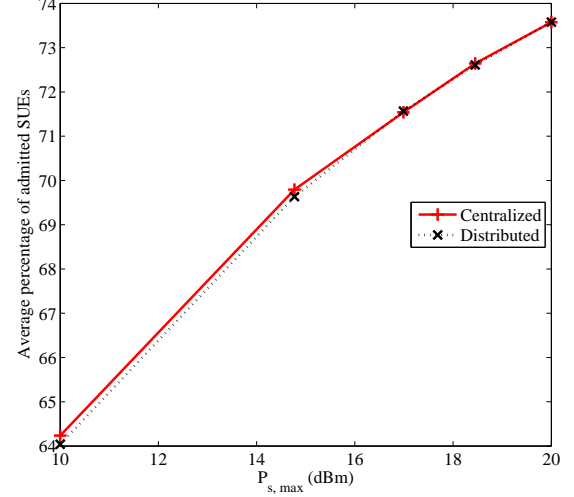
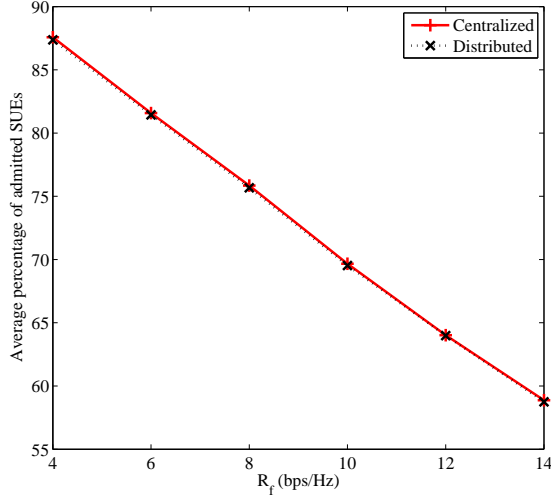
for 50 different realizations, where in each realization, the UE positions and channel gains are varied. Fig. 10 shows the average percentage of admitted SUEs vs. R_f . We notice that, generally, as the rate requirement increases, more SUEs are in outage. We also notice that the distributed scheme converges approximately to the same solution as the centralized solution.

5) *Average percentage of admitted SUEs vs. $P_{s,max}$* : We have the same setup as the one for the previous figure except for $R_f = 10$ bps/Hz. Fig. 11 shows the average percentage of admitted SUEs vs. $P_{s,max}$. We notice that as the maximum transmit power of the small cells increases, the average number of admitted SUEs increases. This rate of increase, however, is not fixed as the system is limited by the interference constraints for MUEs.

C. Summary of Major Observations

The major observations from the numerical analysis can be summarized as follows:

- In a multi-tier network, it is critical to consider the impact of RA decisions in one tier on the other one. For the macrocell network, as different RA schemes are used to achieve the same rate requirements for the MUEs, they affect the performance of the small cell tier differently.

Fig. 9. Convergence of **Algorithm 2**.Fig. 11. Average percentage of admitted SUEs vs. $P_{s,max}$.Fig. 10. Average percentage of admitted SUEs vs. R_f .

- Additional network tiers can be accommodated by minimizing the used bandwidth. The proposed problem formulation for resource allocation in the macrocell leads to a minimal use of the system bandwidth which allows to admit a higher number of SUEs.
- It is foreseen that, employing other schemes, such as a one to “maximize the sum-rate”, for either the macrocell tier or the small cell tier, is not a good option as it will not be able to maximally accommodate additional tiers. Similar to the “minimize the total sum-power” scheme, the “maximize the sum-rate” scheme is known to consume most of the available bandwidth [27].
- For a given macrocell RA policy, increasing the rate requirements for the MUEs degrades the performance of small cell tier in terms of the average number of admitted SUEs.

- By exploiting the time sharing property (i.e., the UEs time share the sub-channels when served by the small cells), a convex optimization formulation can be developed for the RA and AC problem for the small cells. This convex formulation enables us to solve the problem efficiently in a distributed fashion. The distributed algorithm for resource allocation for the small cell tier converges to the same solution as the centralized solution.
- If the deployment of the small cells is such that they are well isolated, resource allocation in the macrocell might have very small effect on the performance of small cells.

VI. CONCLUSION

We have proposed a complete framework for the resource allocation and admission control problem in a two-tier OFDMA cellular network. Different optimization problems with new objectives have been formulated for the macrocell tier and the small cell tier. The macrocell tier aims at allocating resources to its MUEs in a way that can tolerate the maximum possible interference from the small cell tier. This problem has been shown to be an MINLP. However, we have proved that the macrocell can solve another alternate optimization problem that yields the optimal solution in polynomial time. Now, given the interference constraints for the MUEs, the small cells perform resource allocation and admission control with the objective of maximizing the number of admitted SUEs and serving them with the minimum possible bandwidth. This problem has also been shown to be an MINLP. A convex relaxation has been used to study the behavior of the MINLP formulation. Since centralized solutions for resource allocation are not practical for dense networks, a distributed solution for resource allocation and admission control has been proposed using dual decomposition technique and has been shown to converge to the same solution as the centralized one. Numerical results have shown the significance of tier-aware resource allocation methods.

REFERENCES

- [1] A. Abdelnasser and E. Hossain, "Joint resource allocation and admission control in ofdma-based multi-tier cellular networks," in *IEEE Global Communications Conference (GLOBECOM)*, submitted., 2014.
- [2] Ericsson, "5G Radio Access, Research and Vision," *white paper*, 2013.
- [3] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution towards 5G multi-tier cellular wireless networks: An interference management perspective," *arXiv preprint arXiv:1401.5530*, 2014.
- [4] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, pp. 82–89, February 2014.
- [5] D. Lopez-Perez, A. Valcarce, G. de la Roche, and J. Zhang, "OFDMA femtocells: A roadmap on interference avoidance," *IEEE Communications Magazine*, vol. 47, pp. 41–48, September 2009.
- [6] 3GPP, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; (Release 11)," TS 36.300, 3rd Generation Partnership Project (3GPP), Sep 2013.
- [7] K. Son, S. Lee, Y. Yi, and S. Chong, "REFIM: A practical interference management in heterogeneous wireless access networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1260–1272, June 2011.
- [8] V. Nguyen Ha and L. Bao Le, "Fair resource allocation for OFDMA femtocell networks with macrocell protection," *IEEE Transactions on Vehicular Technology*, vol. 63, pp. 1388–1401, March 2014.
- [9] S. Guruacharya, D. Niyato, D. I. Kim, and E. Hossain, "Hierarchical competition for downlink power allocation in OFDMA femtocell networks," *IEEE Transactions on Wireless Communications*, vol. 12, pp. 1543–1553, April 2013.
- [10] D. T. Ngo, S. Khakurel, and T. Le-Ngoc, "Joint subchannel assignment and power allocation for OFDMA femtocell networks," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 342–355, January 2014.
- [11] A. Abdelnasser, E. Hossain, and D. I. Kim, "Clustering and resource allocation for dense femtocells in a two-tier cellular OFDMA network," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 1628–1641, March 2014.
- [12] M. Andersin, Z. Rosberg, and J. Zander, "Gradual removals in cellular pcs with constrained power control and noise," in *IEEE 6th International Symposium on Personal, Indoor and Mobile Radio Communications, 1995. PIMRC'95. Wireless: Merging onto the Information Superhighway*, vol. 1, pp. 56–60 vol.1, 1995.
- [13] D. Lopez-Perez, X. Chu, A. Vasilakos, and H. Claussen, "Power minimization based resource allocation for interference mitigation in OFDMA femtocell networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 333–344, February 2014.
- [14] J. Tadrous, A. Sultan, and M. Nafie, "Admission and power control for spectrum sharing cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1945–1955, 2011.
- [15] I. Mitliagkas, N. Sidiropoulos, and A. Swami, "Joint power and admission control for ad-hoc and cognitive underlay networks: Convex approximation and distributed implementation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 4110–4121, 2011.
- [16] L. B. Le and E. Hossain, "Resource allocation for spectrum underlay in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5306–5315, 2008.
- [17] W. J. Shin, K. Y. Park, D. I. Kim, and J. W. Kwon, "Large-scale joint rate and power allocation algorithm combined with admission control in cognitive radio networks," *Journal of Communications and Networks*, vol. 11, no. 2, pp. 157–165, 2009.
- [18] K. Phan, T. Le-Ngoc, S. Vorobyov, and C. Tellambura, "Power allocation in wireless multi-user relay networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 5, pp. 2535–2545, 2009.
- [19] X. Gong, S. Vorobyov, and C. Tellambura, "Joint bandwidth and power allocation with admission control in wireless multi-user networks with and without relaying," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1801–1813, 2011.
- [20] S. E. Nai, T. Quek, M. Debbah, and A. Huang, "Slow admission and power control for small cell networks via distributed optimization," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2013, pp. 2261–2265, 2013.
- [21] S. Namal, K. Ghahboosi, M. Bennis, A. MacKenzie, and M. Latva-aho, "Joint admission control and interference avoidance in self-organized femtocells," in *Conference Record of the 44th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, 2010, pp. 1067–1071, 2010.
- [22] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser ofdm with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 1747–1758, Oct 1999.
- [23] H. Zhang, W. Zheng, X. Chu, X. Wen, M. Tao, A. Nallanathan, and D. Lopez-Perez, "Joint subchannel and power allocation in interference-limited OFDMA femtocells with heterogeneous QoS guarantee," in *IEEE Global Communications Conference (GLOBECOM)*, 2012, pp. 4572–4577, Dec 2012.
- [24] D. Astely, E. Dahlman, G. Fodor, S. Parkvall, and J. Sachs, "LTE release 12 and beyond [accepted from open call]," *IEEE Communications Magazine*, vol. 51, pp. 154–160, July 2013.
- [25] H. Li, X. Xu, D. Hu, X. Tao, P. Zhang, S. Ci, and H. Tang, "Clustering strategy based on graph method and power control for frequency resource management in femtocell and macrocell overlaid system," *Journal of Communications and Networks*, vol. 13, no. 6, pp. 664–677, 2011.
- [26] W. L. Winston, *Operations Research: Applications and Algorithms*. Cengage Learning, 2003.
- [27] K. Seong, M. Mohseni, and J. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *2006 IEEE International Symposium on Information Theory*, pp. 1394–1398, 2006.
- [28] E. Karipidis, N. Sidiropoulos, and L. Tassiulas, "Joint QoS multicast power / admission control and base station assignment: A geometric programming approach," in *IEEE 5th Sensor Array and Multichannel Signal Processing Workshop, (SAM)*, pp. 155–159, 2008.
- [29] J. Currie and D. I. Wilson, "OPTI: Lowering the Barrier Between Open Source Optimizers and the Industrial MATLAB User," in *Foundations of Computer-Aided Process Operations* (N. Sahinidis and J. Pinto, eds.), (Savannah, Georgia, USA), 8–11 January 2012.
- [30] Z. Shen, J. Andrews, and B. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 2726–2737, Nov 2005.
- [31] M. Tao, Y.-C. Liang, and F. Zhang, "Resource allocation for delay differentiated traffic in multiuser OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 2190–2201, June 2008.
- [32] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.
- [33] S. Boyd, L. Xiao, A. Mutapcic, and J. Mattingley, "Notes on decomposition methods," *Notes for EE364B, Stanford University*, 2007.
- [34] S. Boyd and C. Barratt, "Ellipsoid method," *Notes for EE364B, Stanford University*, 2008.
- [35] N. Komodakis, N. Paragios, and G. Tziritas, "MRF energy minimization and beyond via dual decomposition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 531–552, March 2011.
- [36] 3GPP, "Further advancements for E-UTRA physical layer aspects (Release 9)," TR 36.814, 3rd Generation Partnership Project (3GPP), Mar 2010.